*Risk Analysis, Vol. 32, No. 3, 2012* DOI: 10.1111/j.1539-6924.2011.01692.x

# Equation Reliability of Soil Ingestion Estimates in Mass-Balance Soil Ingestion Studies

**Edward J. Stanek III,[1,*] Bo Xu,[1] and Edward J. Calabrese[2]**

Exposure to chemicals from ingestion of contaminated soil may be an important pathway with potential health consequences for children. A key parameter used in assessing this exposure is the quantity of soil ingested, with estimates based on four short longitudinal mass-balance soil ingestion studies among children. The estimates use trace elements in the soil with low bioavailability that are minimally present in food. Soil ingestion corresponds to the excess trace element amounts excreted, after subtracting trace element amounts ingested from food and medications, expressed as an equivalent quantity of soil. The short duration of mass-balance studies, different concentrations of trace elements in food and soil, and potential for trace elements to be ingested from other nonsoil, nonfood sources contribute to variability and bias in the estimates. We develop a stochastic model for a soil ingestion estimator based on a trace element that accounts for critical features of the mass-balance equation. Using results from four mass-balance soil ingestion studies, we estimate the accuracy of soil ingestion estimators for different trace elements, and identify subjects where the difference between Al and Si estimates is larger (>3RMSE) than expected. Such large differences occur in fewer than 12% of subjects in each of the four studies. We recommend the use of such criteria to flag and exclude subjects from soil ingestion analyses.

**KEY WORDS:** Children; exposure assessment; meta-analysis; mixed models; risk assessment

## 1. INTRODUCTION

An important route of exposure to environmental contaminants, particularly among children, is through ingestion of contaminated soil.[(1)] Exposure is typically estimated by multiplying the concentration of the contaminant in soil by the amount of soil ingested. As a result, an estimate of the amount of soil ingested is needed. In lieu of measuring soil directly, measures are typically made of the quantity of a trace element contained in soil. If the only source of the trace element in a fecal sample is ingested soil, and the trace element is not absorbed, then the quantity of the trace element in feces can be used to estimate the amount of soil ingested. Soil ingestion estimates based on aluminum (Al), silicon (Si), and titanium (Ti) were made using this approach by Binder *et al*.[(2)]

There are several practical problems with this approach. The first problem is that soil is not the only source of trace element ingestion. All trace elements considered have detectable quantities in typical food. In order to estimate the amount of soil ingested, the trace element intake from food must be subtracted from the trace element amount in fecal samples. The trace element amount ingested from food is

[1]Division of Biostatistics and Epidemiology, Department of Public Health, School of Public Health and Health Sciences, University of Massachusetts at Amherst, Amherst, MA, USA.
[2]Division of Environmental Health, Department of Public Health, School of Public Health and Health Sciences, University of Massachusetts at Amherst, Amherst, MA, USA.
*Address correspondence to Edward J. Stanek III, 401 Arnold House, 715 N. Pleasant St., University of Massachusetts, Amherst, MA 01003, USA; stanek@schoolph.umass.edu.

typically estimated from duplicate food samples. A second problem, related to the first, is that it has not been possible to directly distinguish food from soil trace element amounts in fecal samples. As a result, studies have relied on associating food ingestion time periods with fecal output periods, hoping that the corresponding trace element totals will identify the appropriate trace element amounts in food to be subtracted from fecal samples. For example, a 4-day food ingestion period has been associated with a 4-day fecal output period, with the fecal period shifted to be one day later, as in the study by Davis *et al.*[3] The quantity of trace element ingested in food in the 4-day time period is subtracted from the quantity of trace element collected in fecal samples in an effort to eliminate fecal trace element amounts due to food. The success of this strategy depends upon correspondence between the trace element amount in actual food contained in the fecal samples, and the trace element amounts in the collected food samples that are presumed to be contained in the fecal sample.

As the length of a soil ingestion study is extended, the correspondence between the collected duplicate food and the food contained in fecal samples is anticipated to increase. However, for relatively short soil ingestion studies (such as 3- or 4-day studies), there may be appreciable error associated with the linking of food and fecal samples. Furthermore, if soil ingestion is assessed over shorter periods (such as a day), the potential error and resulting variability is increased.

Aspects of a particular trace element can influence reliability. An ideal trace element will be homogeneously distributed in soil, and the coefficient of variation of the trace element concentration in soil will be small, so that the trace element will be able to be reliably measured in fecal samples subsequent to ingestion. At the same time, the concentration of the trace element will be low in food/beverage/medications, so that an error in matching the ingestion period to the fecal sample period will result in a small difference in the soil ingestion estimate.

Biases can occur in several ways. First, trace elements may be ingested from nonfood/nonbeverage/nonmedication/nonsoil sources. Such source errors will positively bias soil ingestion estimates. Since the quantity of trace element from such sources is not known, these biasing effects can only be observed indirectly by comparing trace element amounts measured in fecal samples relative to the expected amount of trace element in the fecal sample based on measured intake and estimated soil ingestion (based on other trace elements). Episodic source error can be best examined over a short time period where it is more likely that a nonfood/soil source occurs for a single trace element. Biases due to ingestion of nonfood/soil sources that contain all trace elements cannot be completely eliminated. Systematic biases due to ingestion of a nonfood source that is not included in the duplicate food collection (such as toothpaste) will positively bias soil ingestion estimates.

Negative bias can occur as a result of bioavailability of a trace element or due to incomplete or missing fecal/urine sample collection. Soil ingestion studies do not enable direct estimates of bioavailability, and such bioavailability, while thought to be minimal, may exist. The extent to which trace elements ingested are not excreted in urine/fecal samples will negatively bias soil ingestion estimates.

Soil ingestion estimates have varied widely between individual studies when based on different trace elements. Although efforts have been made to understand the sources of variability and bias,[4–7] no systematic model has been proposed that accounts for inherent variability in the basic data gathered, and the sources of uncertainty. We develop such a model, introduce stochastic elements to the model, and use the model to assess the reliability of trace element soil ingestion estimates via a simulation study based on mass-balance soil ingestion study data.

We first summarize key study design features of four primary soil ingestion studies that are considered the primary source for childhood soil ingestion estimates in the United States. The studies are described in primary publications by Calabrese *et al.*[8,9,10] and Davis *et al.*[3,11] Using these designs as background, we define a flexible model for mass-balance soil ingestion studies that can be adapted to the individual studies. Using the basic distributions of trace element intake and output from soil ingestion studies, and assuming a distribution of soil ingestion, we evaluate the accuracy of mass-balance soil ingestion estimators in different study designs via simulation studies. Finally, we use estimates of the accuracy and reliability of trace element soil ingestion estimates to illustrate how source errors can be identified in mass-balance soil ingestion studies. We conclude by discussing the application of this approach in the conduct of a meta-analysis of soil ingestion in children.

## 2. THE PRIMARY SOIL INGESTION STUDIES

Our analysis is focused on four mass-balance soil ingestion studies that have been conducted among children in the United States.

### 2.1 Amherst Study

The Amherst study was conducted in late September/October in 1986 and included 65 children between the ages of 1 and 4 years whose parents volunteered to participate in the study. Once enrolled, duplicate food/beverage/medication samples were collected on Monday through Wednesday (3 days) on two consecutive weeks, while total fecal and urine was collected beginning at noon on Monday through noon on Friday (4 days) in the same week. Over-the-counter medications and vitamins were included in duplicate food/beverage samples. Toothpaste was not collected; wipes and toilet paper were not collected. Parents of children in diapers were provided cotton cloth diapers. Parents reported less than 1% of samples were missed or lost. Soil samples were collected from three play areas identified by parents at home and by daycare providers at daycare, and trace element concentrations in soil were weighted to reflect the estimated time a child played in each area at home or at the daycare. Eight trace element amounts (Al, Si, Ti, barium (Ba), manganese (Mn), vanadium (V), yttrium (Y), and zirconium (Zr)) were measured each 24-hour period for intake (pooling food, beverages, and medicine) and output (pooling fecal and urine samples). Soil ingestion estimates were reported on 64 children for each trace element.[8,9,12–15] Subsequent reports discussed variability of estimates between trace elements,[4,5,16–21] very high soil ingestion (pica),[22,23] and trace element ingestion from nonfood/nonsoil sources.[17,24] Daily soil ingestion estimates were given by Stanek and Calabrese.[25]

### 2.2. Washington State Study

The Washington State study[3] was conducted in Pasco, Richland, and Kennewick, Washington in July–September 1987 based on a random digit-dialing sample of households. A total of 138 households were identified with children not in diapers between the ages of 2–7, of which 104 households agreed to participate. Once children were enrolled, duplicate food/beverage samples were collected on 4 consecutive days. Usual medication, dietary supplements, and mouthwash were collected and separately analyzed. Fecal samples were collected on 4 days beginning 1 day after the start of food collection, and approximately half the urine samples were collected over the same time period. Toilet paper was collected, accounting for on average 1.9% of the fecal weight. Families were provided with standard toothpaste previously analyzed for trace elements. Activity time and other behavioral factors were collected during the study period. Soil samples were collected from five locations in play areas and pooled. Samples of food, feces, toilet paper, urine, and fecal samples were pooled over the four study days for each subject, and analyzed for three trace elements (Al, Si, and Ti). Food, urine, and fecal weights were inflated to adjust for reported missing samples. Soil ingestion estimates were reported for each trace element for 101 children by Davis *et al.*[3]

### 2.3. Washington Family Study

A follow-up study, the Washington Family study,[11] was conducted in the summer of 1988 among 20 families identified from participants in the Washington State study. Children were selected who were less than 8 years old at the time of the follow-up study, highly compliant in the initial study, willing to participate, and had two parents/guardians living at home with the child. The protocol for soil ingestion in this study was similar to the Washington State study with the following differences. Children were studied over 13 consecutive days, with food/beverages collected and pooled from days 2–5, and days 6–12, and feces collected and pooled over similar time periods lagged by 1 day. Soil ingestion estimates were made using a mass-balance approach for the 7-day period (days 6–12). Each participant was provided a tube of toothpaste, but toothpaste was not collected. Soil samples were collected from five activity areas identified by an activity study conducted over days 4–7, resulting in a single trace element soil concentration per subject. Samples of food, urine, feces, and toilet paper were pooled over 7 days (days 6–12 for food input, days 7–13 for output) for each subject, and analyzed for three trace elements (Al, Si, and Ti). Food, urine, and fecal weights were inflated to adjust for partial sampling and reported missing samples. Soil ingestion estimates are reported on 12 children for each trace element.

### 2.4. Anaconda Study

The Anaconda study[10] was conducted in Anaconda, Montana in October 1992 on children selected from a stratified simple random sample balanced for gender and age ranging from 1 to 4. Families with eligible children were excluded if there was split custody of the child, if the child attended a daycare or had a disability, or if the family had problems keeping appointments/urine samples. A total of 64 children were selected and enrolled in the study from an initial frame of 258 children. Once enrolled, duplicate food/beverage/medication samples were collected on 7 consecutive days (Monday through Sunday), while total fecal and urine was collected on 7 consecutive days lagging the food collection period by 1 day. Over-the-counter medications and vitamins were included in duplicate food/beverage samples. Toothpaste was provided to all participants that contained nondetectable quantities of trace elements with the exception of silica, which was present in trace quantities. Parents of children in diapers were provided cotton cloth diapers. Wipes and toilet paper were not collected. Parents reported less than 1% of samples were missed or lost. Activity time outdoors was used to identify up to three play areas for pooling soil samples to reflect the estimated time a child played in each area. Eight trace element amounts (Al, Si, Ti, Y, Zr, cerium (Ce), lanthanum (La), and neodymium (Nd)) were measured each 24-hour period for intake (pooling food, beverages, and medicine) and output (pooling fecal and urine samples).

It has proved difficult to summarize soil ingestion results from the different studies. For example, median soil ingestion estimates based on Al have ranged from 25.3 to 36.7 mg/day, while median soil ingestion estimates based on Ti for the same subjects and time period have ranged from 55 mg/day to 206.9 mg/day. Approaches that take advantage of factors that may be related to uncertainty (such as the ratio of the quantity of trace element in food relative to soil) have been used to develop methods of combining different trace element estimates.[4] However, no model has been proposed for soil ingestion that accounts directly for the stochastic features embedded in the mass-balance equation. We describe such a model next, and show how simulations based on this model can be used to identify source biases specific for trace elements due to source error in a mass-balance soil ingestion study.

## 3. A MODEL FOR MASS-BALANCE SOIL INGESTION STUDIES

The soil ingestion studies provide a context for defining a soil ingestion model for trace element intake and output via a mass-balance equation. We develop such a model here under the idealized assumptions that there is no trace element bioavailability, no sample loss, no measurement error, and no missed input or output samples. Trace element intake consists of the trace element amount in food, beverages, medications, and sometimes toothpaste (which we represent by $x$ and refer to as "food") and the trace element amount ingested from soil and/or dust (which we represent by $z$ and refer to as "soil"). Output consists of the trace element amount in fecal and urine, which we refer to as "fecal." The soil ingestion studies measure trace element amounts in intake and output collection periods, which we refer to as "food" and "fecal" collection periods (in days) of duration $\delta_x^*$ and $\delta_y^*$ (where $\delta_x = 24\delta_x^*$ and $\delta_y = 24\delta_y^*$ are corresponding periods in hours), respectively.

The food and fecal collection periods each correspond to a time interval, where the fecal collection period is typically started after the food collection period by some lag, $\lambda$ (see Fig. 1). Although not observed directly, there is a latent intake accumulation period during which the trace element collected in the fecal collection period is actually ingested. The actual time interval corresponding to the intake accumulation period and its duration, $d$, are typically not known. Fig. 1 is idealized since it depicts the intake accumulation period as coinciding with the food collection period, $\delta_x$. In general, these two periods will differ.

We define the transit time as the time taken for the trace element amount ingested in the interval $\Delta t$ (centered at time $t$) to appear in the fecal output. Transit times are not typically observed. For simplicity, we assume that the transit time is discrete. This assumption is a good approximation when the time interval, $\Delta t$, is short. We also assume that the temporal sequence of trace element intake is preserved in trace element output. Two transit times link the intake accumulation period to the fecal collection period. The first transit time, $d_a$, is the time taken for trace elements ingested at the start of the intake accumulation period to reach the start of the fecal collection period. The second transit time, $d_b$, is the time taken for trace elements ingested at the end of the intake accumulation period to reach the end of the fecal collection period.

We consider Fig. 1 to be idealized, since the intake accumulation period exactly matches the food collection period. In such a setting, the transit time $d_a$ is equal to $\lambda$, the lag between the start of the food collection and fecal collection periods. Although the

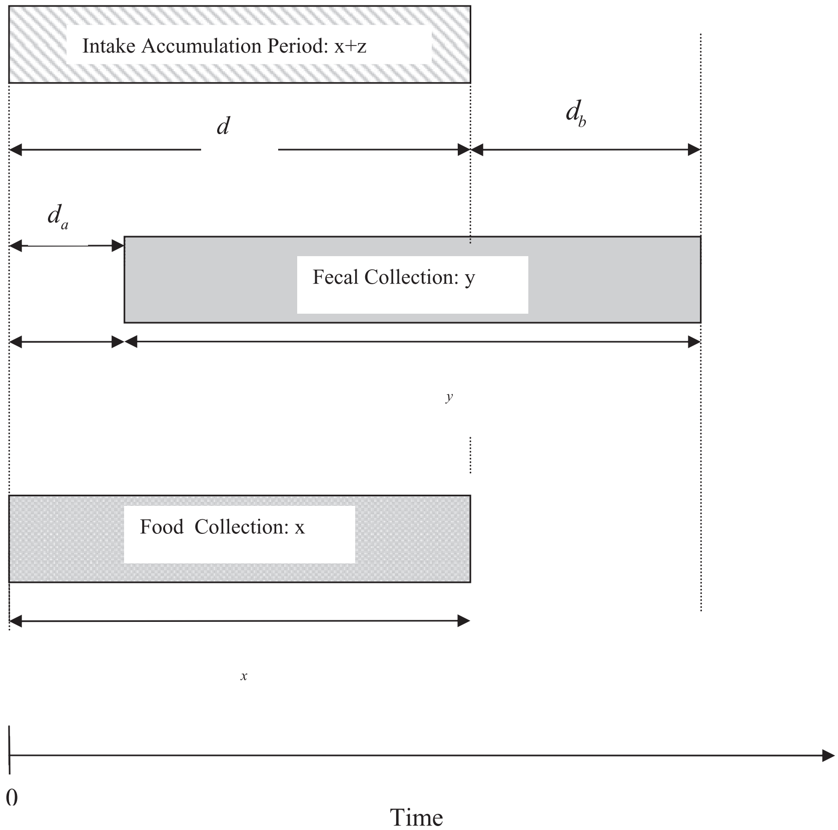



**Fig. 1.** Description of fecal and food collection periods in soil ingestion studies.

food collection period and the fecal collection period are both fixed and known in soil ingestion studies, the intake accumulation period is not known. Different patterns of overlap can occur between the intake accumulation period and the food and fecal collection period. In each of these patterns, the intake accumulation period will always precede the fecal collection period. However, the overlap of the food collection period and the intake accumulation periods may differ. These differences are important in determining the portion of the food collected that actually is contained in the intake accumulation period.

The start of the intake accumulation period, $a_1$, must be prior to the start of the fecal collection period but may be before (patterns 1–4) or after (patterns 5–8) the start of the food collection period, $n_1$ (Fig. 2). The end of the intake accumulation period, $a_2$, may be before the start of the food collection period (pattern 1), during the food collection period (patterns 2, 3, 5, and 6), or after the end of the food collection period, $n_2$ (patterns 4, 7, and 8), but cannot be after the end of the fecal collection period. In all of the soil ingestion studies we consider, the food collection and fecal collection periods overlap, so that by design, pattern 8 does not occur. We assume such an overlap occurs in developing a mass-balance model for soil ingestion.

The principal quantity of interest in soil ingestion studies is the soil ingestion rate, which is defined over the intake accumulation period, and expressed as a rate per day. Let the total trace element amount ingested from soil in the intake accumulation period be represented by $z$. The soil ingestion rate depends on the trace element concentration in the ingested soil, $c_0$, which we assume is constant over the intake accumulation period. The rate of soil ingestion over the intake accumulation period is then given by $\rho = \frac{z}{c_0 d}$, where $d = d_a + \delta_y - d_b$ is the length of the intake accumulation period. Unfortunately, none of the values of $z$, $c_0$, or $d$ are directly observable in mass-balance soil ingestion studies.

For illustration, consider a child in the Anaconda study where duplicate food samples were collected beginning Monday for 7 consecutive days, and fecal samples were collected beginning on Tuesday for 7 consecutive days. There was a 24-hour lag in start times between food and fecal sample collection

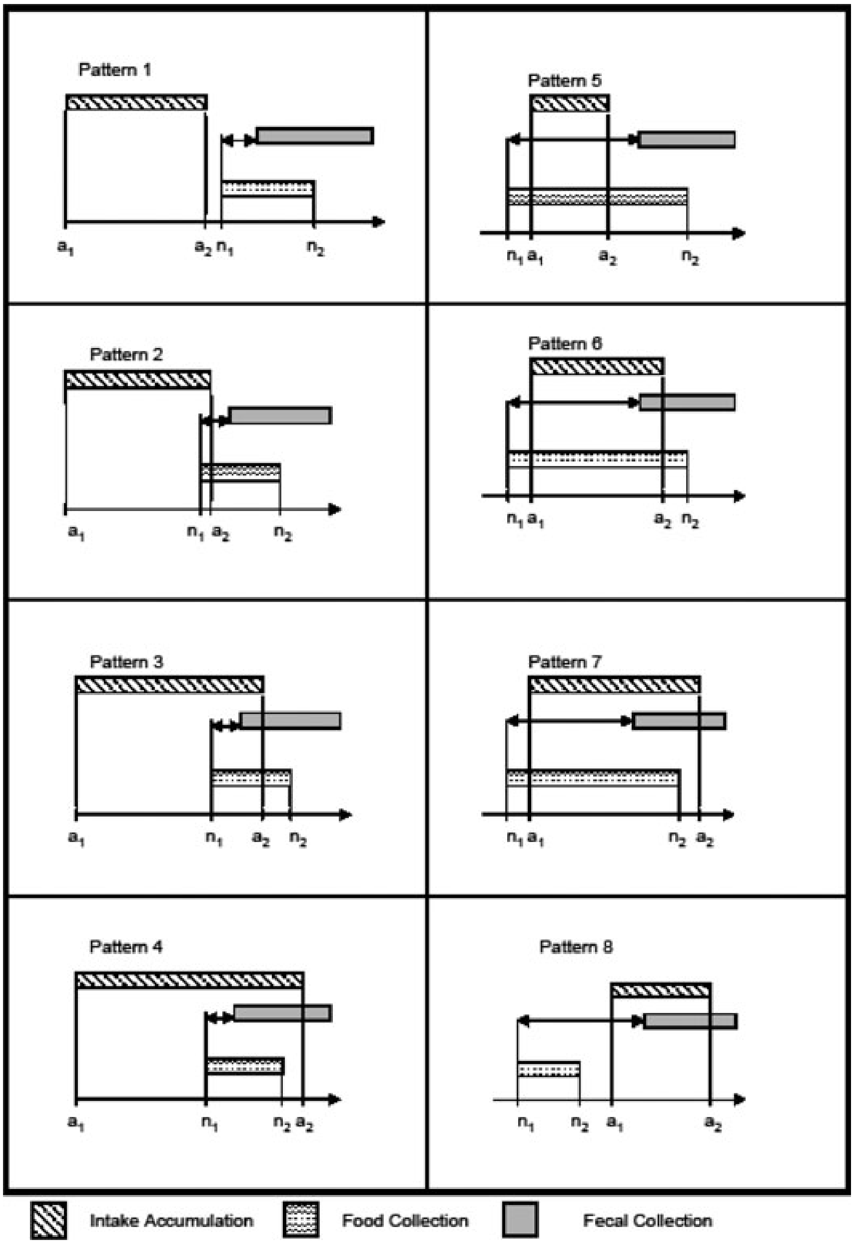



**Fig. 2.** Patterns of overlap of intake accumulation periods with food and fecal collection periods in soil ingestion studies.

periods. This 24-hour lag was set based on an assumption that the transit time from ingestion of food to fecal output is 24 hours. If this assumption holds, the fecal collection period will account for 7 days of intake.

For a child, the intake accumulation period is an important time period that cannot be directly observed, but represents the period when intake occurred for the fecal output. Two transit times define this period. Suppose for a child in the Anaconda study the start of the intake accumulation period was Sunday at noon, and the intake accumulation period ended at noon on the following Saturday. The time it takes for food ingested at noon on Sunday to be excreted in the fecal collection period (starting at midnight on Monday) is 36 hours, which is equal to the transit time, $d_a$. Similarly, food ingested at the end of the intake accumulation period (at noon on Saturday), which was excreted just prior to midnight on Monday took 60 hours to be excreted, which is equal to the transit time, $d_b$. This pattern of transit times can be diagramed as in pattern 3 of Fig. 2. The time points on the abscissa correspond to $a_1$ (noon on Sunday), $n_1$ (midnight on Monday), $a_2$ (noon on the following Saturday), and $n_2$ (midnight on the following Monday). The Anaconda study design included a $\delta_y^* = 7$ day (i.e., $\delta_y = 168$ hour) fecal collection period. Using the transit times for this child, the intake accumulation period is $d = d_a + \delta_y - d_b = 144$ hours, or 6 days. For this child, only 6 days of intake are included in 7 days of fecal collection. The variability in transit times results in variability in the exposure period for soil ingestion from the assumed exposure period of 7 days that would occur if transit time was constant.

Let $x$ represent the total trace element amount ingested from food over the intake accumulation period. Also, let $y$ represent the total trace element intake in the intake accumulation period (equal to the output in the fecal collection period), where $y = z + x$. In practice, $y$ can be observed directly, while neither $x$ nor $z$ can be directly observed. In order to isolate the trace element intake from soil, we use the total trace element intake ingested from food in the food collection period, $x^*$, to estimate the rate of trace element intake from food in the intake accumulation period, $\frac{x^*}{\delta_x}$, and use it to estimate $x$ in the intake accumulation period. We do not observe the length of the intake accumulation period, $d$, but estimate it by the length of the fecal collection period, $d_I$. Finally, using $c$ to estimate the concentration of the trace element in the ingested soil, the soil ingestion estimate is given by:

$$\tilde{r} = \frac{1}{c\delta_y}\left(z + x - \delta_y \frac{x^*}{\delta_x}\right) = \frac{1}{c\delta_y}(z) + \frac{1}{c\delta_y}\left(x - \delta_y \frac{x^*}{\delta_x}\right).$$

The first term in the second line, $z/c\delta_y$, represents an estimate of the average soil ingestion per day (assuming the fecal collection period and intake accumulation period are of equal duration, and using an estimate of the trace element concentration in soil). The second term is an adjustment in the estimate to account for different durations of fecal and food collection periods for trace element intake from food. The soil ingestion estimate differs from $\rho$ for three reasons. First, an estimate of the trace element concentration in soil, $c$, is used in place of the actual trace element concentration, $c_0$. Second, the duration of the fecal collection period $\delta_y$ is used in place of $d_I$, the actual duration of the intake accumulation period, and third, the trace element ingestion rate from food in the food collection period, $\frac{x^*}{\delta_x}$, is used to estimate the trace element ingestion rate for food in the intake accumulation period given by $\frac{x}{d_I}$.

Expressing the soil ingestion estimate as $\tilde{r} = \frac{y}{c\delta_y} - \frac{x^*}{c\delta_x}$, the estimate is the difference in "soil equivalent" rates, where input (from food) is subtracted from output (from food and soil). The sensitivity of this estimate to variability in the soil equivalent rate from food was the basis for earlier approaches to evaluate disparate soil ingestion estimates.[4]

In a mass-balance soil ingestion study, since only $\delta_x$, $\delta_y$, $x^*$, and $y$ are observable, the rate of soil ingestion, $\rho$, and hence the difference, $\tilde{r} - \rho$, cannot be directly observed. However, it is possible to simulate the underlying data generation process, and use the simulation to evaluate the bias and the mean squared error of the soil ingestion estimate $\tilde{r}$ for a given soil ingestion rate. The simulation is based on estimates of the distribution of each random variable, where in many cases, parameters for the distribution are estimated directly from soil ingestion study data. We describe the stochastic model next, and subsequently use it in simulation studies to evaluate the accuracy of soil ingestion estimators for different trace elements. Using soil ingestion estimates from a subset of trace elements that are considered sufficiently accurate, we discuss how these results can be used to screen for potential outliers in mass-balance soil ingestion studies.

### 3.1. Representing Soil Ingestion via a Stochastic Model

We simulate soil ingestion via a stochastic model that represents trace element intake from food and soil, transit time durations, and estimated concentrations of the trace element in soil. The first step is determining the transit times $d_a$ and $d_b$ that connect the intake accumulation period to the fecal collection period. We simulate transit times by assuming that transit times are independent and identically distributed as $D \sim N(\mu_d \;\; \sigma_d^2)$, where $\mu_d = \sigma_d = 24$ hours, truncated so that $D > 0$. Three of the four mass-balance soil ingestion studies were designed to have a 1-day lag between the start of the food collection period and the start of the fecal collection period. We assume that this lag is the mean of the transit time distribution. The standard deviation in transit time was set under the assumption that the mean plus two standard deviations, corresponding to a transit time of 3 days, would be longer than the transit time that actually occurred for 97.5% of the children. The values of $d_a$ and $d_b$, along with the study design variables, identify the response pattern, and more particularly, the regions (Fig. 3) that are defined by the

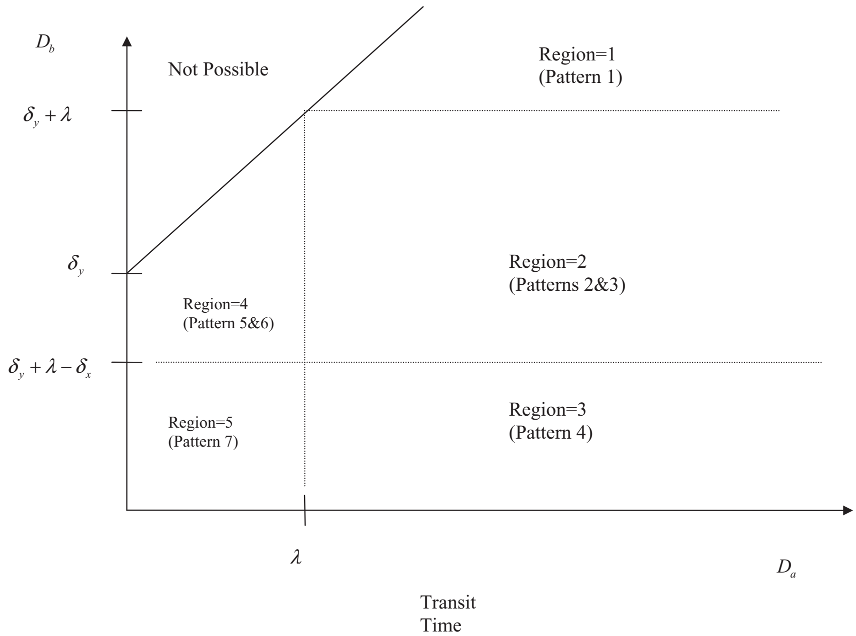



**Fig. 3.** Bivariate transit times regions and corresponding patterns of intake overlap, where $D_a$ and $D_b$ are transit times from the start (and end) of the intake accumulation period to the start (and end) of the fecal collection periods, $\delta_x$ and $\delta_y$ are the durations of the food and fecal collection periods, respectively, and $\lambda$ is the lag between the start of the food collection and fecal collection periods.

**Table I.** Transit Time Regions from the Start of the Intake Accumulation Period to the Start of the Fecal Collection Period, $d_a$, and from the End of the Intake Accumulation Period to the End of the Fecal Collection Period, $d_b$

| Region | Patterns | Range for $d_a$ | Range for $d_b$ |
|---|---|---|---|
| 1 | 1 | $d_a > d_b - \delta_y$ | $\delta_y + \lambda < d_b$ |
| 2 | 2, 3 | $d_a > \lambda$ | $\lambda + \delta_y - \delta_x < d_b < \delta_y + \lambda$ |
| 3 | 4 | $d_a > \lambda$ | $d_b < \lambda + \delta_y - \delta_x$ |
| 4 | 5, 6 | $d_a < \lambda$ | $\lambda + \delta_y - \delta_x < d_b < d_a + \delta_y$ |
| 5 | 7 | $d_a < \lambda$ | $d_b < \lambda + \delta_y - \delta_x$ |

**Table II.** Study Design Characteristics for Amherst, Washington State, and Anaconda/Family Studies

| | Study | | |
|---|---|---|---|
| Characteristic (days) | Amherst | Wash. State | Anaconda/Family |
| Food Duration ($\delta_x$) | 3 | 4 | 7 |
| Fecal Duration ($\delta_y$) | 4 | 4 | 7 |
| Fecal Start Lag ($\lambda$) | 0.5 | 1 | 1 |

ordering of the actual time points for the start and end of the intake accumulation and food collection periods. The range of $d_a$ and $d_b$ for each region is summarized in Table I, with the study design characteristics for the four studies summarized in Table II.

Once the region is determined, three time periods can be defined that correspond to intake of food and/or soil. As an example, suppose that pattern 3 represented the intake and output for a child as illustrated in Fig. 2. For the child, assume that the intake accumulation period corresponded to 10 days, with 2 days overlapping the food collection period. Also assume that the fecal collection period began 1 day after the start of the food collection period, and both the food collection and fecal collection periods are each of 7 days duration. With these assumptions, setting $a_1 = 0$ (corresponding to day 0), the start of the food collection period is $n_1 = 8$ (corresponding to day 8), the start of the fecal collection period is day 9, the end of the intake accumulation period is day $a_2 = 10$, and the end of the food collection period is at the start of day $n_2 = 15$. The three time periods for intake correspond to the period from day 0 to day 8 (prior to the start of the food collection period), the time period from day 8 to day 10 (during the intake accumulation and food collection period), and the period from day 10 to day 15 (after the intake accumulation period but during the food collection period).

This example illustrates how three periods of intake were constructed for each child. Although in the example, each period corresponds to an integer number of days, the simulation of transit times was based on hours, so that the periods may correspond to fractions of days. For a period of length $d_0$ (in hours), we express the period length as an integer number of days, $d_0^*$, where $d_0^* - 1 \leq d_0 \leq$

$d_0^*$. Trace element ingestion from food and soil is simulated on $d_0^*$ consecutive days, with the appropriate proportion of the last day's values used on the last day. We describe the detailed process of simulating trace element ingestion from food and soil next.

Simulation of trace element ingestion from food over the period $d_0$ is based on simulating the freeze-dried food weight and simulating the trace element concentration in food for each day in the period. Data are available from two studies (the Amherst and Anaconda studies) of freeze-dried food weight based on duplicate food samples collected on children. We use these data to estimate the mean freeze-dried food weight, and variance components between subjects, and between days within subjects. Histograms of these data indicate that the freeze-dried weights are approximately normally distributed. Parameters for the mean intake, $\mu_{fdw}$, and variance components for subjects, $\sigma^2_{fdw,S}$, and days, $\sigma^2_{fdw,E}$, are estimated using a mixed model with random subject and day effects. We simulate a subject's latent freeze-food dried weight intake by a random selection from the estimated subject's latent freeze-dried food weight distribution, $N(\hat{\mu}_{fdw}\ \hat{\sigma}^2_{fdw,S})$. Consecutive days' freeze-dried food weight intake for a subject is this latent value, plus a random variable representing day to day variability, selected from the normal distribution, $N(0\ \hat{\sigma}^2_{fdw,E})$, with selections independent between days. In order to account for the portion of the day included at the end of the period, we multiply the last day's intake by $d_0 - d_0^* + 1$, the proportion of the last day included in period $d_0$.

The trace element amount in food on a day depends on the concentration of the trace element in food. Data on trace element concentrations on a daily basis are available from the Amherst and Anaconda studies. The distribution of trace element concentrations was highly skewed to the right, but a log-normal distribution appeared to approximate the reported results. We fit a mixed model to the log-normal concentration distribution data, and estimated the mean concentration, $\mu^*_{X,Tr}$, and variance components for subjects, $\sigma^{*2}_{X,Tr,S}$, and days, $\sigma^{*2}_{X,Tr,E}$. A similar procedure was used to generate the log(normal) trace element concentration on a day for a subject. We exponentiate this concentration to obtain the trace element concentration, and multiplied it by the corresponding freeze-dried food weight to estimate the trace element intake on the day. A similar procedure was used to account for fractions of days at the end of the period.

Daily freeze-dried food weight estimates from the Anaconda and Amherst studies are given in Table III, along with descriptive statistics on the log trace element food concentrations. In order to evaluate the accuracy of mass-balance study soil ingestion estimators, we need to know the true amount of soil ingested. We assume that this amount is known, and specify it by simulating soil ingestion for a day using a log-normal distribution. Five different sets of (mean, std) parameters for the log-normal distribution were selected to provide set of distributions that spanned the range of what is believed to contain the true soil ingestion distribution. The median amount of soil ingested (when averaged for a subject over 4 days or 7 days, corresponding to fecal collection periods in the studies) ranged for the simulated soil ingestion distribution from 19 to 119 mg/day, while the 95% soil ingestion ranged from 79.2 to 841.5 mg/day. For simplicity, we report results for one set of log-normal parameters (mean = 3.5, std = 0.75), but note that the selection of these tracers based on accuracy is not altered by making different assumptions about the soil ingestion distribution.

Daily freeze-dried food weight estimates from the Anaconda and Amherst studies are given in Table III, along with descriptive statistics on the log trace element food concentrations. We assume soil ingestion (mg/day) is log normally distributed on a day for a subject, and define parameters for the mean and variance as 3.5 and 0.5625, respectively, in the log normal distribution to agree approximately with simple results from soil ingestion studies. The assumptions result in a median soil ingestion of 33.1 mg/day; the median plus one standard deviation soil ingestion of 70.1 mg/day; and the median plus two standard deviations of 148.41 mg/day. The assumed parameters in the log-normal distribution are similar to those proposed by Thompson and Burmaster[26] of 4.07 and 0.7225, respectively.

In order to determine the amount of a trace element ingested from soil, we need to know the concentration of the trace element in soil. We assume that the concentration of a trace element in soil is homogenous, and equal to the mean trace element concentration in soil in a given study. This concentration is used to determine the amount of the trace element ingested from soil. Estimates of soil ingestion for subjects have been based on trace element concentrations measured in assumed "play areas" for the subject. We simulate these concentrations by assuming the trace element concentration in soil is normally distributed, with parameters given in

**Table III.** Estimated Parameters for Daily Freeze-Dried Food Weight (g/day) and Natural Logarithm of the Daily Trace Element Food Concentration (mg/g) for Anaconda and Amherst Studies

| | | Subjects Days | | $\mu^*$ | $\sigma^2_{Sx^*}$ | $\sigma^2_{x^*}$ |
|---|---|---|---|---|---|---|
| Characteristic | Study | $n$ | $m$ | Mean | Betw. Subj | Betw. Days |
| Freeze-Dried wt | Anaconda | 64 | 448 | 181.77 | 1541.26 | 2266.19 |
| | Amherst | 64 | 383 | 187.71 | 1387.32 | 1454.16 |
| | Wash. State | 101 | 404 | 251.10 | 4083.24 | 1020.81 |
| | Family | 12 | 84 | 275.28 | 3386.81 | 483.83 |
| Natural Ln Food Conc | | | | | | |
| Al | Anaconda | 64 | 448 | −4.752 | 0.2284 | 0.9288 |
| | Amherst | 64 | 383 | −5.402 | 0.4020 | 1.6973 |
| | Wash. State | 101 | 404 | −3.704 | 0.3454 | 0.0864 |
| | Family | 12 | 84 | −3.959 | 0.4454 | 0.0636 |
| Si | Anaconda | 64 | 448 | −2.6616 | 0.0843 | 0.2744 |
| | Amherst | 64 | 383 | −2.6313 | 0.1160 | 0.2649 |
| | Wash. State | 101 | 404 | −2.4522 | 0.1450 | 0.0363 |
| | Family | 12 | 84 | −3.0947 | 0.2663 | 0.0380 |
| Ti | Anaconda | 64 | 426 | −6.6361 | 1.4882 | 3.3581 |
| | Amherst | 64 | 366 | −7.5725 | 1.0942 | 2.4590 |
| | Wash. State | 101 | 404 | −5.8299 | 1.7968 | 0.4492 |
| | Family | 12 | 84 | −4.6466 | 2.0244 | 0.2892 |
| Y | Anaconda | 64 | 317 | −12.19 | 0.3663 | 0.3389 |
| | Amherst | 64 | 383 | −11.65 | 0.2024 | 0.5604 |
| Zr | Anaconda | 64 | 447 | −9.55 | 0.0373 | 0.1516 |
| | Amherst | 64 | 383 | −10.11 | 0.8413 | 0.3140 |
| Ba | Amherst | 64 | 383 | −6.800 | 0.1070 | 0.1732 |
| Mn | Amherst | 64 | 383 | −5.016 | 0.1862 | 0.1411 |
| V | Amherst | 64 | 383 | −10.18 | 0.1033 | 0.8004 |
| Ce | Anaconda | 64 | 355 | −11.98 | 0.2398 | 0.3163 |
| La | Anaconda | 64 | 259 | −12.50 | 0.8604 | 0.3186 |
| Nd | Anaconda | 64 | 280 | −11.86 | 0.2876 | 0.1814 |

Table IV from the soil ingestion studies. A lower bound equal to 0.1 times the mean concentration of the trace element in soil is used to avoid artificially small concentrations.

With these assumptions, the trace element amount ingested from food and soil in the intake accumulation period is generated, along with the trace element amount ingested in the food collection period. Correlations between these amounts are preserved between trace elements by using a common food freeze-dried weight and a common soil ingestion amount. Correlations between food in the intake accumulation period and the food collection period are preserved by matching appropriate periods. Using the simulated data for a subject, we can evaluate the total quantity of soil ingested, and the soil ingestion rate per day, $\rho$. Notice that the duration of the intake accumulation period varies between subjects due to simulated transit times. Also, using the simulated data, an estimate of soil ingestion can be made corresponding to $\tilde{r}$. We use these estimates to evaluate the bias and variance of soil ingestion estimators for different trace elements in the soil ingestion studies.

## 4. SIMULATION OF SOIL INGESTION

We simulate soil ingestion for 10,000 subjects based on an assumed trace element ingestion distribution for food and soil, with parameters in the distributions estimated from individual soil ingestion studies. For each subject, we compare the soil ingestion rate, $\rho$, with the mass balance estimator, $\tilde{r}$, for each element, evaluating the bias and the variance. We illustrate the distribution of soil ingestion rates for the subjects in Fig. 4.

Subtracting the true soil ingestion rate from the estimated rate, $\tilde{r} - \rho$, we estimate the bias and the standard deviation of this difference for each estimator (Table V). The results indicate minimal bias average bias for all trace element estimators (with the exception of the bias for Ti in the Anaconda study). The

**Table IV.** Estimated Mean and Standard Deviation (mg/g soil) of Trace Element Concentration in Soil for Anaconda and Amherst Studies

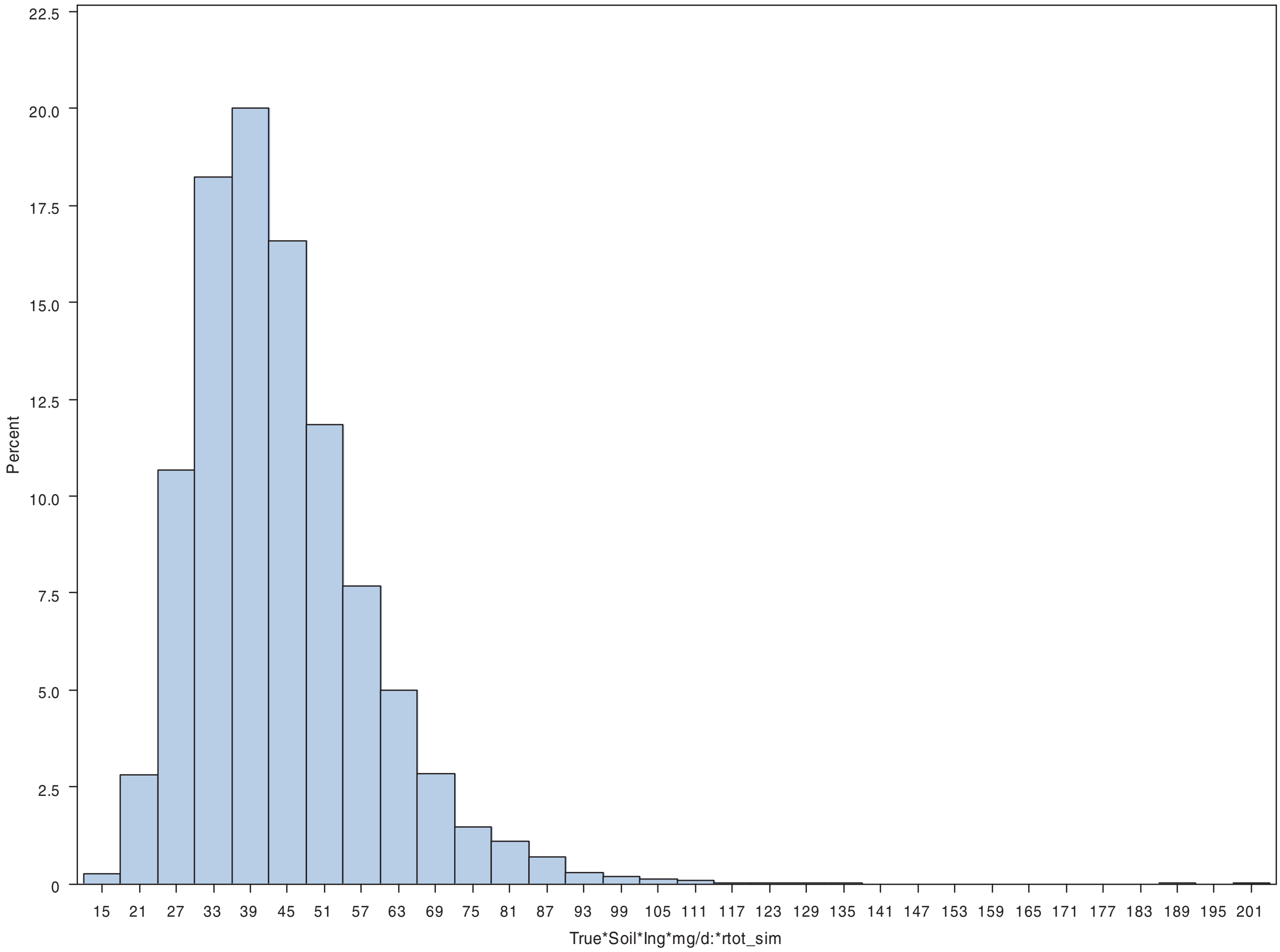
| Characteristic | Study | Subjects *n* | Measures *m* | $\mu^*$ Mean | $\sigma_e$ Betw. Measures |
|---|---|---|---|---|---|
| Al | Anaconda | 64 | 64 | 50.49 | 12.25 |
| Amherst | 64 | 128 | 53.99 | 7.78 | |
| Wash. State | 101 | 101 | 66.10 | 3.84 | |
| Family | 12 | 12 | 65.34 | 1.93 | |
| Si | Anaconda | 64 | 64 | 233.6 | 52.7 |
| Amherst | 64 | 128 | 306.7 | 42.6 | |
| Wash. State | 101 | 101 | 288.9 | 16.3 | |
| Family | 12 | 12 | 277.2 | 6.3 | |
| Ti | Anaconda | 64 | 64 | 1.90 | 0.437 |
| Amherst | 64 | 128 | 3.41 | 0.838 | |
| Wash. State | 101 | 101 | 5.81 | 0.976 | |
| Family | 12 | 12 | 6.09 | 1.069 | |
| Y | Anaconda | 64 | 64 | 0.0157 | 0.0034 |
| Amherst | 64 | 124 | 0.0240 | 0.0037 | |
| Zr | Anaconda | 64 | 64 | 0.1404 | 0.0423 |
| Amherst | 64 | 128 | 0.1959 | 0.0821 | |
| Ba | Amherst | 64 | 128 | 0.3559 | 0.1249 |
| Mn | Amherst | 64 | 128 | 0.7312 | 0.1868 |
| V | Amherst | 64 | 124 | 0.0865 | 0.1108 |
| Ce | Anaconda | 64 | 64 | 0.0210 | 0.0065 |
| La | Anaconda | 64 | 64 | 0.0126 | 0.0041 |
| Nd | Anaconda | 64 | 64 | 0.0097 | 0.0025 |

**Table V.** Summary of Bias, Variance, and RMSE (mg/d) for Trace Elements in Studies

| Characteristic | Study | Bias | Std | RMSE |
|---|---|---|---|---|
| Al | Anaconda | 3.7 | 27.60 | 27.84 |
| Amherst | 0.7 | 47.01 | 47.02 | |
| Wash. State | −1.3 | 52.30 | 52.31 | |
| Family | 0.3 | 26.65 | 26.65 | |
| Si | Anaconda | 2.7 | 25.15 | 25.29 |
| Amherst | 0.7 | 32.09 | 32.09 | |
| Wash. State | −0.2 | 37.56 | 37.56 | |
| Family | 0.0 | 16.54 | 16.54 | |
| Ti | Anaconda | −58.2 | 7934.2 | 7934.4 |
| Amherst | 5.5 | 307.77 | 307.82 | |
| Wash. State | −6.4 | 338.51 | 338.57 | |
| Family | −3.6 | 612.05 | 612.06 | |
| Y | Anaconda | 2.6 | 29.19 | 29.30 |
| Amherst | 1.4 | 56.95 | 56.97 | |
| Zr | Anaconda | 5.5 | 38.99 | 39.37 |
| Amherst | 10.0 | 79.94 | 80.56 | |
| Ba | Amherst | 7.2 | 337.18 | 337.26 |
| Mn | Amherst | −4.4 | 798.97 | 798.98 |
| V | Amherst | −0.1 | 13.39 | 13.40 |
| Ce | Anaconda | 5.9 | 38.85 | 39.30 |
| La | Anaconda | 7.0 | 48.78 | 49.28 |
| Nd | Anaconda | 3.4 | 49.97 | 50.08 |

results also indicate that the soil ingestion estimator based on Ti is highly variable relative to the soil ingestion estimators based on Al or Si. The square root of the MSE (RMSE) is at least six times larger for Ti than Al, and at least nine times larger for Ti than Si in all four soil ingestion studies.

Additional panels in Table V present the results of simulation studies for other trace elements. Two of these trace elements, (Y and Zr) were used in more than one soil ingestion study. The RMSE for these tracers suggest that Y may be of comparable reliability as Al or Si, with Zr somewhat less reliable (with the RMSE at least 1.4 times larger). Among the other trace elements evaluated in a single soil ingestion study, V had the smallest RMSE among all trace elements. The three trace elements Ce, La, and Nd had RMSE at least 1.4 times the RMSE of Al and Si, but still much lower than that of Ti. The trace elements Ba and Mn had high RMSE, comparable to the RMSE of Ti. The results in Table V suggest that soil ingestion estimates based on Ti (and Ba and Mn) are unreliable, and of limited value in estimating soil ingestion.




*Source:* mt10es75.sas on 09/27/2010 by Ed Stanek.

**Fig. 4.** Histogram of simulated soil ingestion mg/d on 10,000 subjects based on a 7-day soil ingestion study design (as in Anaconda).

A principal objective of the simulation studies was the desire to use the simulation study results to characterize variability that we would expect to occur between trace-element-specific estimators for a specific study. The estimates of variability can be used to distinguish normal variability from unusual variability that may be attributable to source error for a particular trace element.

Practically, in order to assess source error, an estimate of soil ingestion that is not thought to be biased is compared with an estimate based on a "test" trace element where source error may have occurred. If the estimate for the test trace element differs substantially (by more than three standard deviations) from the unbiased estimator, source error may be suspected. In order to screen for source error, we need an unbiased estimate of soil ingestion. Initially, we planned on using a weighted least-squares estimator, using one over the MSE as trace element specific weights using all trace elements except the "test" trace element. Review of the results for estimates of soil ingestion (see Table VI) indicated that if a source error occurred for a trace element (such as V) that was considered reliable based on the simulation study results, the source error detection process could fail.

For this reason, a more robust method of estimating soil ingestion for comparison was needed. In addition to using criteria for the MSE to identify reliable trace element soil ingestion estimates, we added criteria to select trace elements that had low potential for source error. Noting that the concentration of trace elements in soil varied between trace elements by over 30,000 fold from Table IV, we hypothesized

**Table VI.** List of WLS Estimate and Trace Element Specific Soil Ingestion Estimates (mg/d) for Five Subjects by Week for the Amherst Study

| ID | Week | V | Si | Al | Y | Zr | Ti | Ba | Mn | WLS |
|---|---|---|---|---|---|---|---|---|---|---|
| 102 | 1 | 8 | 2 | 34 | 20 | 23 | 9 | −668 | 212 | 300 |
| 102 | 2 | 30 | 22 | 59 | 44 | 8 | 9 | 1417 | 358 | 392 |
| 103 | 1 | −111 | −132 | 1 | −54 | −182 | −70 | 155 | −570 | −1361 |
| 103 | 2 | −24 | −60 | 100 | 173 | −4 | −62 | −1058 | 33 | 373 |
| 104 | 1 | −3 | −1 | 76 | −3 | −113 | −114 | −2248 | −893 | −890 |
| 104 | 2 | 26 | 39 | 21 | −27 | −56 | −36 | −117 | −525 | −851 |
| 105 | 1 | 1589 | 2085 | 218 | 14 | 120 | 41 | 188 | −121 | −343 |
| 105 | 2 | 420 | 503 | 249 | 93 | 127 | 29 | 501 | 141 | 384 |
| 107 | 1 | 2095 | 2798 | 45 | 16 | 53 | 10 | −45 | 102 | −318 |
| 107 | 2 | 1737 | 2328 | 11 | 22 | −33 | 24 | 22 | 79 | −201 |

**Table VII.** Percent (and Number) of Measures Greater than 3*RMSE from the Mean (Al/Si) Ingestion Estimate

| Study | Al | Si | Ti | Y | Zr | Ba | Mn | V | Ce | La | Nd |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Anaconda | 6(4) | 6(4) | 0(0) | 27(17) | 8(5) | | | | 28(18) | 34(22) | 67(43) |
| Amherst | 14(18) | 3(4) | 12(15) | 8(10) | 5(6) | 2(3) | 3(4) | 48(62) | | | |
| Wash. State | 14(14) | 14(14) | 15(15) | | | | | | | | |
| Family | 0(0) | 0(0) | 8(1) | | | | | | | | |

that relatively large source error would be less likely to occur for trace elements with large concentrations in soil. The highest concentrations of trace elements in soil are for Al and Si. Also, some of the lowest estimated RMSE occur for Al and Si (Table V). These two considerations led us to estimate the mean soil ingestion for a subject based on a weighted least squares (WLS) estimate using the Al and Si soil ingestion estimates, with weights corresponding to the MSE estimates in Table V. Suspected source error is indicated by a difference in a trace element's estimate from the WLS estimate (based on Al and Si) of greater than three times the RMSE.

The results of this screening for possible source error by trace element are summarized in Table VII. The results indicate the number of subjects (or subject-weeks in the Amherst study) where a soil ingestion estimate for a trace element exceeded 3∗RMSE times the WLS soil ingestion estimate based on Al and Si. The results in Table VII need to be interpreted relative to the reliability results in Table V. For trace elements that are unreliable (Ti, Ba, and Mn), the large RMSE results in a small percentage of the subjects indicated as having a high potential for source error. The lack of reliability for these trace elements makes only extreme source error possible to detect. Among the other trace elements, the trace elements with the least evidence of source error are Al, Si, and Zr. In all studies, source error was indicated in less than 15% of the subjects (/weeks) for these tracers. Other trace elements, such as Y, V, Ce, La, and Nd, had evidence of high levels of source error (over 25%) in nearly all studies. This suggests that casual ingestion of these trace elements from nonfood, nonsoil sources is relatively common, and could bias soil ingestion estimates based on a mass-balance approach. Table VIII lists the WLS soil ingestion estimate as well as the individual estimates for Al and Si that were identified as having possible source error for these elements.

## 5. DISCUSSION

The study is a practical attempt to translate what is known about mass-balance soil ingestion studies into an assessment of the reliability of mass-balance soil ingestion estimators. The problem is complex, since limited observations are available that prevent a close mapping of trace element intake and output. Since studies are conducted on children in a free-living population, and the goal is to estimate typical soil ingestion, the studies have by design lack of control. This means that intake from nonfood, nonsoil sources is possible, and can obscure estimation of the actual soil ingested.

**Table VIII.** List of Subject Periods with Large Differences (>3 RMSE) from WLS Mean for Al and Si

| Study Id | Week | WLS Mean of Al and Si | Soil Ingestion mg/d Al | Soil Ingestion mg/d Si | Rel. Std Diff from WLS Mean |
|---|---|---|---|---|---|
| Amherst 105 | 1 | 153 | 14 | 218 | 3.6 |
| Amherst 118 | 2 | 290 | 11 | 420 | 7.2 |
| Amherst 225 | 1 | 121 | −117 | 232 | 6.1 |
| Amherst 228 | 1 | 218 | 435 | 117 | 5.6 |
| Amherst 228 | 2 | 155 | 287 | 94 | 3.4 |
| Amherst 230 | 2 | 179 | 451 | 52 | 7.0 |
| Amherst 831 | 2 | 149 | 19 | 209 | 3.3 |
| Amherst 832 | 1 | 198 | 55 | 265 | 3.7 |
| Amherst 832 | 2 | 142 | 20 | 199 | 3.2 |
| Amherst 835 | 1 | 199 | 40 | 273 | 4.1 |
| Amherst 843 | 1 | 290 | 51 | 401 | 6.2 |
| Amherst 843 | 2 | 118 | −2 | 174 | 3.1 |
| Amherst 851 | 2 | 11796 | 13600 | 10956 | 46.5 |
| Amherst 854 | 1 | −70 | 145 | −169 | 5.5 |
| Amherst 856 | 1 | 582 | 25 | 842 | 14.4 |
| Amherst 856 | 2 | 219 | 39 | 303 | 4.6 |
| Amherst 862 | 2 | 260 | 427 | 183 | 4.3 |
| Amherst 863 | 2 | 207 | 70 | 271 | 3.5 |
| Anaconda 4 | 52 | −203 | 262 | 12.4 | |
| Anaconda 19 | | 66 | 179 | −28 | 5.5 |
| Anaconda 27 | | 220 | 402 | 69 | 8.9 |
| Anaconda 145 | | 250 | 461 | 75 | 10.3 |
| WashState 109 | | 557 | 873 | 394 | 7.4 |
| WashState 117 | | 313 | 164 | 390 | 3.5 |
| WashState 130 | | 129 | −5 | 199 | 3.2 |
| WashState 150 | | 55 | 199 | −19 | 3.4 |
| WashState 157 | | 262 | 764 | 4 | 11.8 |
| WashState 162 | | 294 | −39 | 465 | 7.8 |
| WashState 173 | | 330 | 35 | 483 | 7.0 |
| WashState 176 | | 130 | −1 | 198 | 3.1 |
| WashState 184 | | 162 | 10 | 240 | 3.6 |
| WashState 199 | | 148 | −24 | 236 | 4.0 |
| WashState 210 | | −247 | 55 | −403 | 7.1 |
| WashState 216 | | 137 | 5 | 205 | 3.1 |
| WashState 218 | | 266 | 103 | 350 | 3.8 |
| WashState 220 | | −55 | −219 | 30 | 3.9 |

The simulation approach provides a controlled way of assessing reliability of soil ingestion estimates in an ideal situation. We attempted to design the simulation so that it closely reflected what we believe to occur in practice. The model for transit times, and their connection to the intake accumulation period, creates a mechanical process that links input to study design characteristics and output. We have accounted for the correlation of trace element intake due to volume, and approximated the distribution of trace element concentrations in food and soil based on estimates from individual soil ingestion studies. We have allowed for variability in the concentration of trace elements in soil by allowing the estimated soil concentration to differ from the average concentration, as has been observed in studies. Simulations have been conducted using five different assumptions concerning daily soil ingestion distributions. The median and 95th percentiles of these distributions span the range of percentiles for soil ingestion thought to occur among children.

Assumptions were made in the simulation study that could alter the results. First, we assume in the simulation that nonfood trace element intake is due to soil, not dust, and do not account for intake from other sources. Similarly, we assumed that trace elements from food and soil proceed in a linear fashion through the digestive system, and that there is no mixing. Mixing certainly occurs, as, for example, with microclustering of large size undigested particles, but its impact has not been incorporated. We have not accounted for "meals," correlation of intake between adjacent meals due to leftover food, or sleep in the simulation. Each may result in higher variability due to microclustering of intake. We also have not accounted for systematic factors such as toothpaste, which may or may not be included in trace element intake. In some studies, a bias from not accounting for toothpaste may occur. We have assumed that food and fecal sample collection is complete, even though in practice some missing collections are reported. We also have assumed that there is no trace element bioavailability, or if bioavailable, that the trace element is not sequestered, but excreted in urine and collected. Finally, we have assumed transit time parameters, and a distribution, but acknowledge that other transit time parameters and transit time distributions may affect the results. Each of these assumptions tends to result in higher estimates of the reliability of soil ingestion. Most likely, the reliability of soil ingestion estimates is less than that reported here. If the day-to-day variability in true soil ingestion is larger than that assumed here, the true reliability of soil ingestion estimates will be further reduced.

Although there are many caveats to the simulation study, we believe that the simulation results provide a scientific basis for identifying reliable tracers, and constructing soil ingestion estimates using a mass-balance approach. The simulation highlights the dependence of mass-balance studies on the transit time, and the uncertainty of the intake accumulation period. This has implication for interpreting soil ingestion estimates, since although soil ingestion studies have been designed over a fixed period, the estimates for different subjects are averages over different time intervals. The simulations illustrate that many of the suggestions of reasons for discrepancies in discussions of disparate soil ingestion estimates actually occur. Variability in food intake will result in soil ingestion estimates of low reliability, especially when accompanied by low concentrations of trace elements in soil. The simulation study provides an objective way to assess such reliability.

Finally, the estimates of reliability from the simulations provide a measure by which possible source error can be assessed for individual trace elements. In past evaluations of soil ingestion data, such error has been postulated for certain trace elements, but a systematic approach has been difficult to develop without introducing subjective criteria. The simulation study estimates of trace element reliability go a long way to eliminating this subjectivity. Coupled with estimates of average soil ingestion based on Al and Si (with high concentrations in soil), subjects where trace elements have potential source error can be identified. This may lead to a more objective framework for pooling data to estimate soil ingestion in children, and ultimately more reliable soil ingestion estimates. Since soil ingestion is one of the most sensitive determinants of dose and risk when assessing soil cleanup criteria for dioxins and other contaminants,[27] this may lead to a sounder scientific basis for determining cleanup levels that protect the public's health.

## ACKNOWLEDGMENTS

We thank Edmund Crouch, Jackie Moya, Bryce Landenberger, Robert Budinsky, and Jeff Crum for helpful comments; Scott Davis and Dana Mirack for the Washington State Study data; and Guosho Yuan, Shuli Yu, and Ruitao Zhang for website construction. The website http://www.umass.edu/metasoil/index.html contains data for each study used in the analysis. This work was supported by a contract to the University of Massachusetts from Dow Chemical Company, Michigan.

## REFERENCES

1. U.S. EPA. Child-Specific Exposure Factors Handbook (Final Report). Washington, DC: U.S. Environmental Protection Agency, EPA/600/R-06/096F, 2008.
2. Binder S, Sokal D, Maughan D. Estimating soil ingestion: The use of tracer elements in estimating the amount of soil ingested by young children. Archives in Environmental Health, 1986; 41:341–345.
3. Davis S, Waller P, Buschbom R, Ballou J, White P. Quantitative estimates of soil ingestion in normal children between the ages of 2 and 7 years: Population-based estimates using aluminum, silicon, and titatium as soil tracer elements. Archives of Environmental Health, 1990; 45:112–122.
4. Stanek EJ, Calabrese EJ. A guide to interpreting soil ingestion studies: I. Development of a model to estimate the soil ingestion detection level of soil ingestion studies. Regulatory Toxicology, 1991; 13:263–277.
5. Calabrese EJ, Stanek EJ III. A guide to interpreting soil ingestion studies: II. Quantitative and qualitative evidence of soil ingestion. Regulatory Toxicology, 1991; 13:278–292.
6. Stanek EJ III, Calabrese EJ, Zorn M. Biasing factors for simple soil ingestion estimates in mass balance studies of soil ingestion. Human and Ecological Risk Assessment, 2001; 7:329–355.
7. Stanek EJ III, Calabrese EJ, Zorn M. Soil ingestion distributions for Monte Carlo risk assessment in children. Human and Ecological Risk Assessment, 2001; 7:357–368.
8. Calabrese EJ, Barnes R, Stanek EJ III, Pastides H, Gilbert CE, Veneman P. How much soil do young children ingest:

An epidemiologic study. Regulatory Toxicology, 1989; 10:123–137.

9. Calabrese EJ, Pastides H, Barnes R, Edwards C, Kostecki PT, Stanek EJ, Veneman P, Gilbert CE. How much soil do young children ingest: An epidemiologic study. Pp. 363–397 in Calabrese EJ, Kostecki PT (eds). Petroleum Contaminated Soils. Boca Raton, FL: Lewis Publishing, 1989.
10. Calabrese EJ, Stanek EJ III, Pekow P, Barnes R. Soil ingestion estimates for children residing on a superfund site. Ecotoxicology and Environmental Safety, 1997; 36:258–268.
11. Davis S, Mirick DK. Soil ingestion in children and adults in the same family. Journal of Experimental Science and Environmental Epidemiology, 2006; 16:63–75.
12. Calabrese EJ, Gilbert C, Kostecki P, Barnes R, Stanek EJ III, Veneman P, Pastides H, Edwards C. Epidemiological study to estimate how much soil children eat. Pp. 313–320 in Calabrese EJ, Kostecki PT (eds). Petroleum Contaminated Soils. Remediation Techniques, Environmental Fate, Risk Assessment. Boca Raton, FL: Lewis Publishing, 1989.
13. Wang X, Lasztity A, Viczian M, Israel Y, Barnes R. Inductively coupled plasma spectrometry in the study of childhood soil ingestion. Part 1. Methodology. Journal of Analytical Atomic Spectro. 1989; 4:727–735.
14. Lasztity A, Wang X, Viczian M, Israel Y, Barnes R. Inductively coupled plasma spectrometry in the study of childhood soil ingestion. Part 2. Recovery. Journal of Analytical Atomic Spectrometry, 1989; 4:737–742.
15. Barnes RM. Childhood soil ingestion: How much dirt do kids eat? Analytical Chemistry, 1990; 62:1–8.
16. Calabrese EJ, Stanek EJ III. Soil ingestion issues and recommendations. Journal of Environmental Science and Health, 1994; A29:517–530.
17. Calabrese EJ, Stanek EJ III. Resolving inter-tracer inconsistencies in soil ingestion estimation. Environmental Health Perspectives, 1995; 103:454–457.
18. Stanek EJ III, Calabrese EJ. Methodological considerations in assessing soil ingestion. Chemical Spectrometry and Bioavailability, 1991; 3:65–68.
19. Stanek EJ III, Calabrese EJ. Bias and the detection limit model for soil ingestion. Journal of Soil Contamination, 1994; 3:183–189.
20. Stanek EJ III, Calabrese EJ. Soil ingestion estimates for use in site evaluations based on the best tracer method. Human and Ecological Risk Assessment, 1995; 1:133–156.
21. Stanek EJ III, Calabrese E, Barnes RM, Keegan E, Lasztity A, Wang X, Gilbert C, Pastides H, Kostecki PT. Ingestion of trace elements from food among preschool children: Al, Ba, Mn, Si, Ti, V, Y, and Zr. Journal of Trace Elements and Experimental Medicine, 1988; 1, 179–190.
22. Calabrese EJ, Stanek EJ III, Gilbert CE. Lead exposure in a soil pica child. Journal of Environmental Science and Health, 1993; A28:353–362.
23. Calabrese EJ, Stanek EJ III. Soil pica: Not a rare event. Journal of Environmental Science and Health, 1993; A28;373–384.
24. Calabrese EJ, Stanek EJ III. High levels of exposure to vanadium by children ages 1–4. Journal of Environmental Science and Health, 1993; A28:2359–2371.
25. Stanek EJ III, Calabrese EJ. Daily estimates of soil ingestion in children. Environmental Health Perspectives, 1995; 103:276–285.
26. Thompson KM, Burmaster DE. Parametric distributions for soil ingestion by children. Risk Analysis, 1991; 2:339–342.
27. Paustenbach DJ, Fehling K, Scott P, Harris M, Kerger BD. Identifying soil cleanup criteria for dioxins in urban residential soils: How have 20 years of research and risk assessment experience affected the analysis? Journal of Toxicology and Environmental Health. Part B, Critical Reviews, 2006; 9(2):87–145.